Annals of Actuarial Science, Vol. 7, part 1, pp. 26–45.
doi:10.1017/S1748499512000280
First published online 19 November 2012

Diversification in heavy-tailed portfolios: properties and pitfalls

Georg Mainik*
RiskLab, Department of Mathematics, ETH Zurich

Paul Embrechts†
RiskLab, Department of Mathematics, ETH Zurich

Abstract

We discuss risk diversification in multivariate regularly varying models and provide explicit formulas for Value-at-Risk asymptotics in this case. These results allow us to study the influence of the portfolio weights, the overall loss severity, and the tail dependence structure on large portfolio losses. We outline sufficient conditions for the sub- and superadditivity of the asymptotic portfolio risk in multivariate regularly varying models and discuss the case when these conditions are not satisfied. We provide several examples to illustrate the resulting variety of diversification effects and the crucial impact of the tail dependence structure in infinite mean models. These examples show that infinite means in multivariate regularly varying models do not necessarily imply negative diversification effects. This implication is true if there is no loss-gain compensation in the tails, but not in general. Depending on the loss-gain compensation, asymptotic portfolio risk can be subadditive, superadditive, or neither.

Keywords

Diversification; Value-at-Risk; heavy tails; tail dependence; risk subadditivity; risk superadditivity; multivariate regular variation; infinite means

1 Introduction

Imagine you manage a portfolio of assets or insurance risks. Building a mathematical portfolio model leads to the problem of adding random variables with general dependence structures. In particular, dependence of risks or asset returns in benign conditions can be quite different from that caused by market crashes or other catastrophic events.

The mathematical basis of portfolio diversification was given by Markowitz (1952) for multivariate Gaussian models. The mean-variance approach of Markowitz can also be extended to the class of elliptical distributions. Under these assumptions, diversification is always good in the sense that the portfolio risk is always minimized by a mixed portfolio. However, this is not true in general. Negative diversification effects in heavy-tailed models with independent α-stable asset returns are known at least since Fama & Miller (1972, Chapter 5).

*Correspondence to: Georg Mainik, RiskLab, Department of Mathematics, ETH Zurich, Raemistrasse 101, 8092 Zurich, Switzerland. E-mail: georg.mainik@math.ethz.ch
†Senior SFI Professor

Various financial asset returns and insurance loss data exhibit heavy tails. One of the earliest references is Mandelbrot (1963), followed by Fama (1965) and many others. A recent empirical study of Moscadelli (2004) concludes that operational risk data may even suggest distributions with infinite means. Another example is to be found in the realm of nuclear losses, see Hofert & Wüthrich (2011). In such cases, portfolio diversification needs tools that go far beyond the classical mean-variance method.

Although the Markowitz approach still remains popular, the recurring financial crises and natural catastrophes raise questions about tail dependence and diversification of extremes. Recent contributions in this area include Rootzén & Klüppelberg (1999), Ibragimov *et al.* (2009), Ibragimov *et al.* (2011), Embrechts *et al.* (2009), Zhou (2010), Mainik & Rüschendorf (2010) and references therein. This list is, of course, by no means complete.

Another mathematical field related to diversification problems is the theory of risk measures. It addresses the general question for mathematical methods to quantify risk. Risk measurement in asymmetric models needs downside risk measures, such as the Value-at-Risk (VaR) and the Expected Shortfall (ES). The question for compatibility of risk measurement with diversification initiated the development of the theory of coherent risk measurement by Artzner *et al.* (1999).

Moreover, there is a recent result in this area that is related to negative diversification effects. As shown by Delbaen (2009), it is impossible to define a coherent risk measure on L^p spaces with $p \in (0, 1)$. The practical consequence of this result is that in case of potentially infinite loss expectations one always has to be aware that diversification can increase the portfolio risk.

One more topic worth mentioning here is the non-coherence of VaR. It is well known that VaR is not subadditive in general. As VaR is very popular in practice, both sufficient conditions for VaR subadditivity and the analysis of models that do not satisfy them are very important.

This paper is dedicated to asymptotic diversification effects in portfolios with heavy-tailed returns. The central model assumption we make is multivariate regular variation ($\mathcal{MRV}$). It implies that all assets are heavy-tailed with the same tail index, and that there is a non-degenerate tail dependence structure. Under this assumption all assets can contribute to the extremal behaviour of the portfolio loss.

$\mathcal{MRV}$ is, of course, a strong simplification of the real world, where tail indices can be different. However, many popular models are $\mathcal{MRV}$. This includes the multivariate Student-*t* distribution, multivariate stable models, models with appropriate copulas, e.g., of Gumbel or Galambos type, and identical, regularly varying margins. Assuming that all assets have relevance to extremes, we reduce the study to the case with non-trivial asymptotics. The results we present give insight into the properties of many models and highlight some pitfalls in this area.

Our main contribution is the answer to the question whether infinite means in $\mathcal{MRV}$ models always imply asymptotic superadditivity of VaR_λ for $\lambda \nearrow 1$. We show that the answer depends on the loss-gain compensation in the tails. On the one hand, this implication is true if large losses cannot be compensated by large gains. This result is related to the non-diversification result from Mainik and Rüschendorf, 2010. On the other hand, we show that this implication is wrong in general. The possibility of loss-gain compensation allows for a wide range of diversification effects. The final result is crucially influenced by the particular tail dependence structure, and VaR may be superadditive, subadditive, or neither. We demonstrate this by examples for each case and discuss some related modelling traps.

One of the models we discuss is used by Daníelsson *et al.* (2005), Daníelsson *et al.* (2012) in simulation studies on VaR subadditivity. The results of this simulation study have not yet been fully explained and may cause confusion if regarded without proper analysis of the underlying model. This model is a particularly interesting example for the influence of the tail dependence structure and the tail index on the diversification effects.

The paper is organized as follows. In Section 2 we introduce the basic notation, the modelling framework of multivariate regular variation, and give an outline of general results on portfolio VaR asymptotics in $\mathcal{MRV}$ models. This includes calculation of the asymptotic portfolio risk and sufficient criteria for sub- or superadditivity. Section 3 comprises examples illustrating these results. Here we discuss the variety of diversification effects in infinite mean models with loss-gain compensation and highlight related modelling traps. In Section 4 we outline model specific and more general conclusions. The spectral measure of the model used by Daníelsson *et al.* (2005), Daníelsson *et al.* (2012) is derived in Appendix A.

2 Portfolio losses and multivariate regular variation

2.1 Basic notation

Consider a random vector $X = (X^{(1)}, \ldots, X^{(d)})$ in $\mathbb{R}^d$ representing risks or asset returns. Focusing on the risky side, let positive component values $X^{(i)}$ represent losses and let the gains be indicated by negative $X^{(i)}$. Then the portfolio loss is given by

$$\xi^\top X := \sum_{i=1}^{d} \xi^{(i)} X^{(i)},$$

where $\xi = (\xi^{(1)}, \ldots, \xi^{(d)})$ is a vector of portfolio weights. If not mentioned otherwise, all vectors in $\mathbb{R}^d$ will be regarded as column vectors. In particular, $\xi = (\xi^{(1)}, \ldots, \xi^{(d)})$ actually means $\xi = (\xi^{(1)}, \ldots, \xi^{(d)})^\top$. To keep the writing as simple as possible, the transposition operation $(\,\cdot\,)^\top$ will be mentioned explicitly only if really necessary.

According to the intuition of diversifying a unit capital over different assets, we restrict ξ to the hyperplane

$$H_1 := \left\{ x \in \mathbb{R}^d : x^{(1)} + \ldots + x^{(d)} = 1 \right\}.$$

Additional constraints on the portfolio weights can be implemented by restriction of ξ to smaller subsets of H_1. A particularly important special case is the exclusion of negative portfolio weights, so-called short positions. The corresponding portfolio set is the unit simplex:

$$\Sigma^d := \left\{ x \in \mathbb{R}^d_+ : x^{(1)} + \ldots + x^{(d)} = 1 \right\}.$$

2.2 Multivariate regular variation

Aiming at dependence of extremes, we assume that the probability distribution of X features a non-trivial dependence structure in the tails. This assumption is made precise in the following definition.

Definition 2.1 A random vector X in $\mathbb{R}^d$ is *multivariate regularly varying* ($\mathcal{MRV}$), if, as $t \to \infty$,

$$\mathcal{L}\big((t^{-1}\|X\|, \|X\|^{-1}X) \mid \|X\| > t\big) \xrightarrow{w} \rho_\alpha \otimes \Psi \tag{1}$$

where ρ_α is the Pareto(α) distribution, i.e., $\rho_\alpha((x, \infty)) := x^{-\alpha}$ for $x \geq 1$, and Ψ is a probability measure on the $\|\cdot\|$-unit sphere $\mathbb{S}^d_{\|\cdot\|} := \left\{ s \in \mathbb{R}^d : \|s\| = 1 \right\}$.

The parameter $\alpha > 0$ is called *tail index* of X, and the measure Ψ in (1) is called *spectral* or *angular measure* of X. In the sequel we will use the short notation $X \in \mathcal{MRV}_{-\alpha,\Psi}$ for multivariate regular variation with tail index α and spectral measure Ψ.

It is obvious that $X \in \mathcal{MRV}_{-\alpha,\Psi}$ entails univariate regular variation of $\|X\|$ with same tail index α. That is, the distribution function $F_{\|X\|}$ of $\|X\|$ satisfies

$$\forall r > 0 \quad \lim_{t\to\infty} \frac{1 - F_{\|X\|}(tr)}{1 - F_{\|X\|}(t)} = r^{-\alpha}. \tag{2}$$

We will also use the short writing $\|X\| \in \mathcal{RV}_{-\alpha}$ for this property. It is well known that the tail index α separates finite moments from the infinite ones: (2) implies that $\mathrm{E}\|X\|^\beta < \infty$ for $\beta < \alpha$ and $\mathrm{E}\|X\|^\beta = \infty$ for $\beta > \alpha$. Regular variation of random variables taking positive and negative values can be considered for the upper and the lower tail separately.

Moreover, $X \in \mathcal{MRV}_{-\alpha,\Psi}$ implies $|X^{(i)}| \in \mathcal{RV}_{-\alpha}$ for all i if

$$\forall i \in \{1, \ldots, d\} \quad \Psi\left(\left\{ s \in \mathbb{S}^d_{\|\cdot\|} : s^{(i)} = 0 \right\}\right) < 1. \tag{3}$$

This non-degeneracy condition guarantees that all components $X^{(i)}$ are relevant to the extremes of the portfolio loss $\xi^\top X$. It should also be noted that $X \in \mathcal{MRV}_{-\alpha,\Psi}$ implies $\xi^\top X \in \mathcal{RV}_{-\alpha}$ for all ξ under appropriate non-degeneracy conditions in the spirit of (3). For further details and for inverse results of Cramér-Wold type we refer to Basrak *et al.* (2002) and Boman & Lindskog (2009).

The $\mathcal{MRV}$ property can also be defined without polar coordinates. One can start with the assumption that there exists a sequence $a_n \to \infty$ and a (non-zero) Radon measure ν on the Borel σ-field $\mathcal{B}([-\infty, \infty]^d \backslash \{0\})$ such that $\nu([-\infty, \infty]^d \backslash \mathbb{R}^d) = 0$ and, as $n \to \infty$,

$$n\mathrm{P}^{a_n^{-1}X} \xrightarrow{v} \nu \text{ on } \mathcal{B}\Big([-\infty, \infty]^d \backslash \{0\}\Big), \tag{4}$$

where $\xrightarrow{v}$ denotes the *vague convergence* of Radon measures and $\mathrm{P}^{a_n^{-1}X}$ is the probability distribution of $a_n^{-1}X$. (cf. Resnick, 2007). This formulation is more technical than (1), but the measure ν in (4) is a very useful object. It is unique except for a multiplicative factor and exhibits the scaling property

$$\forall t > 0 \quad \nu(tA) = t^{-\alpha}\nu(A), \tag{5}$$

which is the key to the most applications of $\mathcal{MRV}$ models. It already implies that $\nu = (c\bar{\rho}_\alpha \otimes \Psi) \circ \tau$ where $\tau(x) := (\|x\|, \|x\|^{-1}x)$ is the polar coordinate transform, $\bar{\rho}_\alpha((x, \infty]) := x^{-\alpha}$ is an extension of ρ_α to $(0, \infty]$, and $c > 0$ is a constant. One can always obtain $c = 1$ by choosing $a_n = F^{\leftarrow}_{\|X\|}(1 - 1/n)$ where $F^{\leftarrow}_{\|X\|}$ is the quantile function of $\|X\|$. We assume this standardization of ν throughout the following.

The measure ν also provides a link to the multivariate Extreme Value Theory. If

$$\nu\left(\left\{x \in \mathbb{R}^d : x^{(i)} > 1\right\}\right) > 0, \quad i = 1, \ldots, d, \tag{6}$$

then ν also characterizes the asymptotic distribution of the componentwise maxima $M_n := (M^{(1)}, \ldots, M^{(d)})$ with $M^{(i)} := \max\{X_1^{(i)}, \ldots, X_n^{(i)}\}$. An equivalent writing of (6) is $\Psi(\{s \in \mathbb{S}^d_{\|\cdot\|} : s_i > 0\}) > 0$ for all i, which is a special case of (3). This assumption implies that

$$\mathrm{P}\{a_n^{-1} M_n \in [-\infty, x]\} \xrightarrow{\mathrm{w}} \exp\left(-\nu\left([-\infty, \infty]^d \setminus [-\infty, x]\right)\right) \tag{7}$$

for $x \in (0, \infty]^d$ and $a_n = F^{\leftarrow}_{\|X\|}(1 - 1/n)$. Therefore ν is also called *exponent measure*. For further details on the asymptotic distributions of maxima see Resnick (1987) and Haan & Ferreira (2006).

Although the domain of Ψ depends on the norm underlying the polar coordinates, the $\mathcal{MRV}$ property is norm-independent in the following sense: if (1) holds for some norm $\|\cdot\|$, then it holds also for any other norm $\|\cdot\|_\diamond$ that is equivalent to $\|\cdot\|$. In this paper we use the sum norm $\|x\|_1 := \sum_{i=1}^d |x^{(i)}|$ and let Ψ denote the spectral measure on the unit sphere $\mathbb{S}^d_1$ induced by $\|\cdot\|_1$. In the special case of $\mathbb{R}^d_+$-valued random vectors it may be convenient to reduce the domain of Ψ to $\mathbb{S}^d_{\|\cdot\|} \cap \mathbb{R}^d_+$. For $\mathbb{S}^d_1$ this is the unit simplex Σ^d.

Further details on regular variation of functions or random variables can be found in Bingham *et al.* (1987), Resnick (1987), Basrak *et al.* (2002), Mikosch (2003), Hult & Lindskog (2006), Haan & Ferreira (2006), Resnick (2007).

2.3 Portfolio loss asymptotics

The $\mathcal{MRV}$ assumption has strong consequences on the asymptotic behaviour of large portfolio losses. It allows to assess the asymptotics of the *Value-at-Risk* VaR_λ and the *Expected Shortfall* ES_λ for $\lambda \nearrow 1$, i.e., far out in the tail. The next result provides a general characterization of the asymptotic portfolio losses in multivariate regularly varying models. The special case of random vectors in $\mathbb{R}^d_+$ was studied in Mainik & Rüschendorf (2010). The general case is treated in Mainik (2010, Lemma 3.2).

Lemma 2.2. *Let $X \in \mathcal{MRV}_{-\alpha,\Psi}$. Then*

(a)

$$\lim_{t\to\infty} \frac{\mathrm{P}\{\xi^\top X > t\}}{\mathrm{P}\{\|X\|_1 > t\}} = \gamma_\xi := \int_{\mathbb{S}^d_1} (\xi^\top s)_+^\alpha \, \mathrm{d}\Psi(s); \tag{8}$$

(b)

$$\lim_{u\nearrow 1} \frac{F^{\leftarrow}_{\xi^\top X}(u)}{F^{\leftarrow}_{\|X\|_1}(u)} = \gamma_\xi^{1/\alpha}. \tag{9}$$

The immediate consequence of (8) and (9) is that the functional γ_ξ characterizes the asymptotics of portfolio loss probabilities and the corresponding high loss quantiles. In particular, the limit relation (9) allows for an asymptotic comparison of the Value-at-Risk associated with different portfolio vectors ξ. The Value-at-Risk $\mathrm{VaR}_\lambda(Y)$ of a random loss Y is defined as the λ-quantile of Y (cf. McNeil *et al.* 2005):

$$\mathrm{VaR}_\lambda(Y) := F_Y^{\leftarrow}(\lambda). \tag{10}$$

In the context of diversification effects, the basic question is the comparison of portfolio losses $\xi_1^\top X$ and $\xi_2^\top X$ for standardized portfolio vectors $\xi_1, \xi_2 \in H_1$. From (10) we immediately obtain an asymptotic comparison of the portfolio VaR.

Corollary 2.3. *Let $X \in \mathcal{MRV}_{-\alpha,\Psi}$ and $\xi_1, \xi_2 \in H_1$. Then*

$$\lim_{\lambda \nearrow 1} \frac{\mathrm{VaR}_\lambda(\xi_1^\top X)}{\mathrm{VaR}_\lambda(\xi_2^\top X)} = \left(\frac{\gamma_{\xi_1}}{\gamma_{\xi_2}}\right)^{1/\alpha}. \tag{11}$$

Analogous comparison results for the Expected Shortfall ES_λ and other *spectral risk measures* are also possible (cf. Mainik & Rüschendorf, 2010).

A particularly important case is $\xi_2 = e_i$, where e_i is the i-th unit vector for $i \in \{1, \dots, d\}$. This portfolio vector represents the single asset strategy investing all capital in the i-th asset. The ratio γ_ξ/γ_{e_i} is well defined if $\gamma_{e_i} > 0$. If $\gamma_{e_i} = 0$ for some i, then the risk optimal portfolio cannot contain any asset j with $\gamma_{e_j} > 0$. The non-degeneracy assumption (6) is equivalent to $\gamma_{e_i} > 0$ for all i. It keeps the losses of all assets on the same scale and focuses the discussion on the non-trivial cases. Henceforth we assume (6) to be satisfied.

2.4 Sub- and superadditivity

The limit relation (11) links asymptotic subadditivity of VaR to the functional $\gamma_\xi^{1/\alpha}$. Indeed, we can write $X^{(i)} = e_i^\top X$ and $X^{(1)} + X^{(2)} = 2\eta^\top X$ with $\eta := \left(\frac{1}{2}, \frac{1}{2}\right)$. Applying (11), we obtain that

$$\frac{\mathrm{VaR}_\lambda(X^{(1)} + X^{(2)})}{\mathrm{VaR}_\lambda(X^{(1)}) + \mathrm{VaR}_\lambda(X^{(2)})} \to \frac{\gamma_{\left(\frac{1}{2},\frac{1}{2}\right)}^{1/\alpha}}{\frac{1}{2}\gamma_{e_1}^{1/\alpha} + \frac{1}{2}\gamma_{e_2}^{1/\alpha}}.$$

Thus we see that checking the asymptotic subadditivity of VaR for $X = (X^{(1)}, X^{(2)})$ is related to the comparison of $\gamma_\eta^{1/\alpha}$ and $\eta^{(1)}\gamma_{e_1}^{1/\alpha} + \eta^{(2)}\gamma_{e_2}^{1/\alpha}$.

A more general approach is the analysis of the mapping $\xi \mapsto \rho(\xi^\top X)$ for some risk measure ρ. Extending from the equally weighted portfolio η to $\xi \in \Sigma^d$, we see that the asymptotic subadditivity of VaR_λ for $\lambda \nearrow 1$ is related to the inequality

$$\gamma_\xi^{1/\alpha} \leq \sum_{i=1}^d \xi^{(i)} \gamma_{e_i}^{1/\alpha}.$$

That is, we need to know whether the mapping $\xi \mapsto \gamma_\xi^{1/\alpha}$ is convex on the unit simplex Σ^d. In terms of diversification, convexity of $\gamma_\xi^{1/\alpha}$ means that a mixed portfolio is typically better than a one-asset strategy.

The convexity of $\gamma_\xi^{1/\alpha}$ is related to the Minkowski inequality on the function space $L^\alpha(\Psi)$. This was already pointed out by Embrechts *et al.* (2009) in a slightly different setting related to the aggregation of risks. Applying (8), one immediately obtains that

$$\gamma_\xi^{1/\alpha} = \|h_\xi\|_{\Psi,\alpha} \tag{12}$$

with $h_\xi(s) := (\xi^\top s)_+$ and $\|f\|_{\Psi,\alpha} := (\int f^\alpha \mathrm{d}\Psi)^{1/\alpha}$. Although $\|\cdot\|_{\Psi,\alpha}$ is not a norm for $\alpha < 1$, one still can define function spaces $L^\alpha(\Psi)$ as collections of all measurable functions $f : \mathbb{S}_1^d \to \mathbb{R}$ such that $\|f\|_{\Psi,\alpha} < \infty$. We demonstrate below that the missing triangle inequality for $\|\cdot\|_{\Psi,\alpha}$ is the origin of asymptotic risk superadditivity in $\mathcal{MRV}$ models with $\alpha < 1$.

The next theorem summarizes the diversification properties in the special case $\Psi(\Sigma^d) = 1$, as it occurs for random vectors in $\mathbb{R}_+^d$. For $\mathbb{R}^d$-valued random vectors, $\Psi(\Sigma^d) = 1$ means that the excess behaviour of the gains is weaker than that of the losses, so that compensation of high losses by high gains is impossible. This setting is typical for risk aggregation in insurance and reinsurance, with small incremental premia constantly coming in and potentially large losses from rare events. In financial applications, risk aggregation without loss-gain compensation is particularly important in the area of operational risk.

Theorem 2.4. *Let* $\mathrm{X} \in \mathcal{MRV}_{-\alpha,\Psi}$ *with* $\alpha > 0$ *and* $\Psi(\Sigma^\mathrm{d}) = 1$, *and restrict the portfolio vector* ξ to Σ^d. *Then the mapping* $\xi \mapsto \gamma_\xi^{1/\alpha}$ *is*

(a) convex for $\alpha > 1$;

(b) linear for $\alpha = 1$;

(c) concave for $\alpha < 1$.

Proof. Let $\xi_1, \xi_2 \in \Sigma^d$ and $\lambda \in (0, 1)$. Then

$$\forall s \in \Sigma^d \quad h_{\lambda\xi_1+(1-\lambda)\xi_2}(s) = \lambda h_{\xi_1} + (1-\lambda) h_{\xi_2}.$$

Thus the case $\alpha = 1$ is trivial and the rest follows from the Minkowski inequality for L^p spaces with $p \in (0, \infty)$. The standard case $p \geq 1$ is well known, whereas for $p < 1$ and non-negative functions the inequality is inverse. (cf. Hardy *et al*. 1934, Theorem 2.24, p.30). □

Remark 2.5 The concavity or convexity in Theorem 2.4 for $\alpha \neq 1$ is strict if Ψ is not concentrated on a linear subspace in the sense that $\Psi(\{s : a^\top s = 0\}) = 1$ for some $a \in \mathbb{R}^d$. This follows from the fact that $\|\lambda h_{\xi_1}\|_{\Psi,\alpha} + \|(1-\lambda)h_{\xi_2}\|_{\Psi,\alpha} = \|\lambda h_{\xi_1} + (1-\lambda)h_{\xi_2}\|_{\Psi,\alpha}$ for $\alpha \neq 1$ implies $h_{\xi_1} = b h_{\xi_2}$ Ψ-a.s. for some $b \geq 0$. (cf. Hardy *et al*. 1934, Theorem 2.24).

There are two general conclusions from Theorem 2.4. On the one hand, if the $\mathcal{MRV}$ assumption accords with the real world data and α is greater than 1, then one can expect VaR_λ to be subadditive for λ close to 1. Although appropriate choice of dependence structure for given marginal distributions always allows to violate the subadditivity of VaR (cf. McNeil *et al*. 2005, Example 6.22, and Embrechts & Puccetti, 2010), $\mathcal{MRV}$ excludes these pathological cases at least in the asymptotic sense. On the other hand, if $\alpha < 1$ and the $\mathcal{MRV}$ assumption fits the reality, then diversification is generally bad for any asymptotic dependence structure Ψ on Σ^d. Surprising as it may appear at the first glance, this phenomenon has an intuitive explanation. The mathematical background of diversification is the Law of Large Numbers, which essentially means that the fluctuation of averages is lower than that of separate random variables. If the expectations are infinite, this reasoning breaks down. In the insurance context, this means that sharing catastrophic risks may increase the danger of insolvency. In the context of operational risk data having a tail index below 1, the increased financial power of a larger bank may still be insufficient to compensate the increased intensity of operational losses. With infinite means in the risk data, one can only reduce the total risk by reducing the number of risk exposures.

The next theorem summarizes the diversification properties in the general case, where Ψ is not concentrated on Σ^d. This setting allows the gains to be on the same scale with losses, so that loss-gain compensation can take place in the tail region.

Theorem 2.6. *Let $X \in \mathcal{MRV}_{-\alpha,\Psi}$, $\alpha > 0$. Then the mapping $\xi \mapsto \gamma_\xi^{1/\alpha}$ is*

(a) continuous;

(b) convex for $\alpha \geq 1$.

Proof. Part (a). As $s \in \Sigma^d$ is bounded, the mapping $\xi \mapsto h_\xi(s)$ is continuous uniformly in $s \in \Sigma^d$. This implies the continuity of the mapping $\xi \mapsto \|h_\xi\|_{\Psi,\alpha}$.

Part (b). Let $\xi_1, \xi_2 \in \mathbb{S}_1^d$ and $\lambda \in (0,1)$. The convexity of the mapping $t \mapsto t_+$ yields

$$\forall s \in \Sigma^d \quad h_{\lambda\xi_1+(1-\lambda)\xi_2}(s) \leq \lambda h_{\xi_1}(s) + (1-\lambda)h_{\xi_2}(s).$$

The result follows from the Minkowski inequality. □

Compared to Theorem 2.4, the most important difference in Theorem 2.6 is the missing statement for $\alpha < 1$. This is not only because the techniques used before do not apply here. Diversification effects for $\alpha < 1$ in models with loss-gain compensation are much more complex than in the pure loss setting. The crucial factor here is the tail dependence structure, i.e., the spectral measure Ψ. For some Ψ one can have convexity, for some others piecewise concavity. Models that appear similar at a first glance turn out to have very different VaR asymptotics. Some of these modelling traps and the resulting confusion will be discussed in the next section.

3 Examples and discussion

3.1 Catastrophic risks: dependence vs. independence

According to Theorem 2.4, superadditivity of portfolio risks is inevitable if $\alpha < 1$ and the losses cannot be compensated by gains. In addition to that, it turns out that the influence of dependence on the diversification effects is inverse in this case. That is, lower dependence of risk components $X^{(i)}$ increases the risk of any portfolio $\xi^\top X$ for $\xi \in \Sigma^d$. The upper bound for asymptotic portfolio risk is attained by the random vector with independent components, whereas the lower bound is achieved by taking all risk components equal. This inverse ordering of diversification effects for $\alpha < 1$ was shown by Mainik & Rüschendorf (2012).

Figure 1 shows the diversification effects arising in a bivariate regularly varying model with a Gumbel copula C_ϑ and identically distributed, non-negative, regularly varying margins $X^{(i)}$. The dependence parameter ϑ ranges from 1 to ∞, thus covering both extremal cases: the independence ($\vartheta = 1$) and the monotonicity ($\vartheta = \infty$). To make the diversification effect curves comparable, the portfolio risk functional is normalized according to (11). That is, the plots show the asymptotic VaR ratio of the portfolio $\xi^\top X$ and the single asset $X^{(1)}$:

$$\gamma_\xi^{*1/\alpha} := \left(\frac{\gamma_\xi}{\gamma_{e_1}}\right)^{1/\alpha} = \lim_{\lambda \nearrow 1} \frac{\mathrm{VaR}_\lambda(\xi^\top X)}{\mathrm{VaR}_\lambda(X^{(1)})}.$$

The ordering of the asymptotic risk profiles $\gamma_\xi^{*1/\alpha}$ with respect to the dependence parameter ϑ suggests a uniform ordering of diversification effects for $\xi \in \Sigma^d$. The direction of this ordering

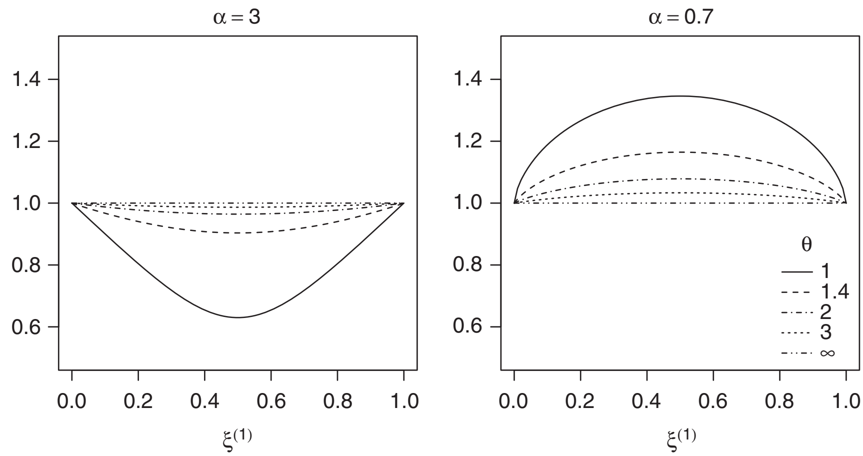

Figure 1. The asymptotic VaR ratio $\gamma_{\xi}^{*1/\alpha}$ for $\mathcal{MRV}$ models with a Gumbel copula: diversification is bad for $\alpha < 1$.

depends on α: smaller ϑ improves the diversification effects for $\alpha > 1$, but increases the portfolio risk for $\alpha < 1$. For the calculation of γ_{ξ}^{*} in this model and a mathematical proof of the ordering result see Mainik & Rüschendorf (2012).

The inverse diversification effects for $\alpha < 1$ and the inverse impact of dependence on the portfolio risk illustrated in Figure 1 are typical for regularly varying models in $\mathbb{R}_{+}^{d}$. Moreover, the parameter values $\vartheta = 1$ and $\vartheta = \infty$ represent ultimate bounds for diversification effects that can be attained at any dependence structure (cf. Mainik & Rüschendorf, 2012). In particular, additive VaR is the best case one can have for $\alpha < 1$.

3.2 Elliptical distributions

An important class of stochastic models is that of elliptical distributions. It can be considered as a generalization of the multivariate Gaussian distribution that preserves the elliptical shape of sample clouds but allows for non-Gaussian tails. In particular, the standard variance-covariance aggregation rules for VaR remain valid for all elliptical distributions (cf. McNeil *et al.* 2005, Theorem 6.8).

A random vector X in $\mathbb{R}^d$ is elliptically distributed if it satisfies

$$X \overset{\mathrm{d}}{=} \mu + RAU,$$

where $\mu \in \mathbb{R}^d$, $A \in \mathbb{R}^{d \times d}$, U is uniformly distributed on the Euclidean sphere $\mathbb{S}_2^d$, and R is a non-negative random variable that is independent of U. If $\mathrm{E}R < \infty$, then $\mathrm{E}X = \mu$, and if $\mathrm{E}R^2 < \infty$, then the covariance matrix of X is given by

$$\mathrm{E}\big((X-\mu)(X-\mu)^{\top}\big) = \mathrm{E}\big(R^2 AUU^{\top}A^{\top}\big) = \mathrm{E}R^2 AA^{\top}.$$

The matrix $C := AA^{\top}$ is called *ellipticity matrix* of X. It is unique except for a constant factor. Given a symmetric and positive semidefinite matrix C, we can always find A such that $C = AA^{\top}$ by Cholesky decomposition.

For elliptically distributed X, $X \in \mathcal{MRV}$ is equivalent to $R \in \mathcal{RV}_{-\alpha}$ for some $\alpha > 0$ (cf. Hult & Lindskog, 2002). To exclude degenerate cases, we assume throughout the following that C is positive definite. The spectral measure of X depends on C and α. Explicit formulas for the spectral density in the bivariate case are derived in Hult & Lindskog (2002), and a general representation for $d \geq 2$ is given in Mainik (2010).

However, the calculation of the asymptotic risk profile $\gamma_\xi^{*1/\alpha}$ for elliptical distributions can be carried out without spectral measures. Let $a := \xi^\top A$. Then

$$\xi^\top X \stackrel{\mathrm{d}}{=} \xi^\top \mu + \|a\|_2 R(\|a\|_2^{-1} a)^\top U.$$

By symmetry of $\mathbb{S}_2^d$ we have that $(\|a\|_2^{-1}a)^\top U \stackrel{\mathrm{d}}{=} e_1^\top U = U^{(1)}$. This gives

$$\xi^\top X \stackrel{\mathrm{d}}{=} \xi^\top \mu + \|\xi^\top A\|_2 Z$$

with $Z \stackrel{\mathrm{d}}{=} R U^{(1)}$.

Hence $\mathrm{VaR}_\lambda(\xi^\top X) = \xi^\top \mu + \|\xi^\top A\|_2 F_Z^{\leftarrow}(\lambda)$. As $F_Z^{\leftarrow}(\lambda) \to \infty$ for $\lambda \nearrow 1$, we obtain from (11) that

$$\gamma_\xi^{*1/\alpha} = \left(\frac{\gamma_\xi}{\gamma_{e_1}}\right)^{1/\alpha} = \lim_{t\to\infty} \frac{\xi^\top \mu + t\|\xi^\top A\|_2}{\mu^{(1)} + t\|e_1^\top A\|_2} = \sqrt{\xi^\top C \xi / C_{1,1}}.$$

It is also easy to see that this diversification effect is non-asymptotic for all centred elliptical distributions with any R (not necessarily regularly varying). That is, for $\mu = 0$ and $\lambda \in \left(\frac{1}{2}, 1\right)$ we always have

$$\frac{\mathrm{VaR}_\lambda(\xi^\top X)}{\mathrm{VaR}_\lambda(X^{(1)})} = \sqrt{\xi^\top C \xi / C_{1,1}}.$$

This is exactly the variance-covariance VaR aggregation rule, which was originally derived in the Gaussian setting.

Figure 2 shows plots of this ratio in the bivariate setting with $C = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ for different values of ρ. In particular, the asymptotic diversification effect $\gamma_\xi^{*1/\alpha} = (\gamma_\xi / \gamma_{e_1})^{1/\alpha}$ does not depend on α if X is elliptical. Moreover, we have a uniform ordering of portfolio risks in the sense that lower ρ implies lower portfolio risk for any $\xi \in \Sigma^d$. This remarkable property is a consequence of the

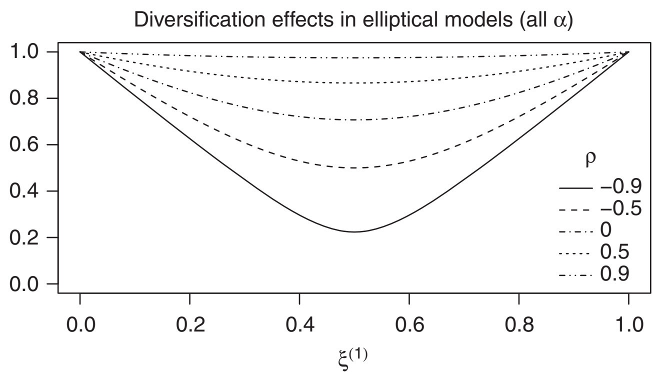

Figure 2. The asymptotic VaR ratio $\gamma_\xi^{*1/\alpha}$ for elliptical distributions does not depend on α.

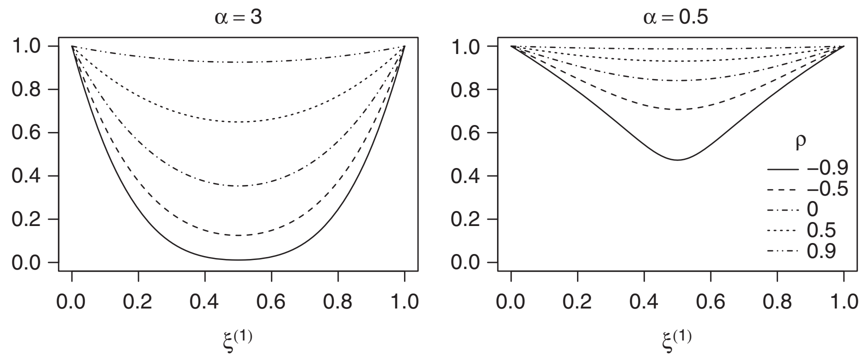

Figure 3. Asymptotic excess probabilities may be misleading: $\gamma_\xi/\gamma e_1$ for elliptical distributions depends on α (whereas $\gamma_\xi^{*1/\alpha} = (\gamma_\xi/\gamma_{e_1})^{1/\alpha}$ does not).

geometric structure of elliptical distributions. On the other hand, the asymptotic ratio of excess probabilities for different portfolios depends on α. Applying (8), we immediately obtain that

$$\lim_{t\to\infty} \frac{\mathrm{P}(\xi^\top X > t)}{\mathrm{P}(X^{(1)} > t)} = \frac{\gamma_\xi}{\gamma_{e_1}} = \left(\xi^\top C\xi/C_{1,1}\right)^{\alpha/2}.$$

Plots of this ratio are shown in Figure 3.

3.3 Heavy-tailed linear models are not elliptical

An appealing property of multivariate Gaussian models is their interpretation in terms of linear regression. In the bivariate case this essentially means that a bivariate Gaussian random vector $X = (X^{(1)}, X^{(2)})$ with margins $X^{(1)}, X^{(2)} \sim \mathcal{N}(0,1)$ and correlation $\rho \in (-1, 1)$ satisfies

$$X^{(2)} = \rho X^{(1)} + \sqrt{1-\rho^2}\, Y^{(2)},$$

where $Y^{(2)} \sim \mathcal{N}(0,1)$, independent of $X^{(1)}$. Setting $Y^{(1)} := X^{(1)}$ and $Y := (Y^{(1)}, Y^{(2)})^\top$, we can write it as

$$X = AY, \quad A := \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}. \tag{13}$$

That is, X can be obtained as a linear transformation of a random vector with independent margins. The generation of $X \sim \mathcal{N}(0, C)$ for an arbitrary covariance matrix $C \in \mathbb{R}^{d\times d}$ uses (13). It suffices to choose A such that $AA^\top = C$. As $\mathcal{N}(0, C)$ is elliptical with ellipticity matrix C, the diversification effects in the model (13) are the same as in Figure 2.

However, the Gaussian case is the only one where the linear model (13) with independent $Y^{(i)}$ is elliptical. To demonstrate the difference between elliptical and linear models in the heavy-tailed case, we compare the multivariate elliptical Student-t distribution with the model generated according to (13) from Y with independent t-distributed margins. The same heavy-tailed linear model was used by Daníelsson *et al.* (2005), Daníelsson *et al.* (2012) in simulation studies on risk sub- and superadditivity. The simulation results obtained there deviated strongly from what one would expect in an elliptical model. In particular, the VaR subadditivity depended on the tail index α, which should not be the case for an elliptical t distribution. The analysis presented below will explain the simulation results of Daníelsson *et al.* (2005) and give additional insight into the behaviour of VaR in $\mathcal{MRV}$ models with $\alpha < 1$.

Let $X = AY$ with A from (13) and assume that the margins $Y^{(1)}, Y^{(2)}$ of Y are independent Student-t distributed with degrees of freedom equal to $\alpha > 0$. Note that $\mathrm{E}(Y^{(i)2}) < \infty$ for $\alpha > 2$. In this case, the correlation matrix of X is well defined and given by

$$\mathrm{Cor}(X) = AA^\top = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

The generation of elliptical t random vectors is quite easy. Let $W^{(1)}, W^{(2)}$, V be independent random variables with $W^{(i)} \sim \mathcal{N}(0,1)$ and $V \sim \chi^2$ with α degrees of freedom. Then the random vector

$$Z := \sqrt{\frac{\alpha}{V}} AW, \quad W = \left(W^{(1)}, W^{(2)}\right)^\top \tag{14}$$

is elliptically distributed with $Z^{(i)} \sim t(\alpha)$, $i = 1, 2$ (cf. McNeil *et al.* 2005, Example 3.7). The ellipticity matrix is equal to C, which is also $\mathrm{Cor}(Z)$ for $\alpha > 2$.

Figure 4 shows scatterplots of 1000 simulated i.i.d. observations of Z and X, respectively. Although Z and X have the same "correlation structure" (correlation is only defined for $\alpha > 2$), the difference between the samples is remarkable. While large observations of Z concord with the elliptical shape of the sample cloud, the excess points of X are concentrated on two axes. This concentration gets stronger for heavier tails, i.e., for smaller α. The cross-shaped sample clouds indicate that the spectral measure of the linear heavy-tailed model (13) consists of four atoms. The calculation of this spectral measure is given in Appendix A.

This property of the linear model originates from the polynomial tails of the t distribution. It is well known that the $t(\alpha)$ distribution is regularly varying with tail index α. Moreover, symmetry arguments give $|Y^{(i)}| \in \mathcal{RV}_{-\alpha}$, i.e., $\bar{F}_{|Y^{(i)}|}(r) = r^{-\alpha} l(r)$ with some $l \in \mathcal{RV}_0$.

The asymptotic diversification effect $\gamma_\xi^{*1/\alpha}$ can be calculated directly. It is obvious that $e_1^\top X = Y^{(1)}$ and

$$\xi^\top X = (\xi^{(1)} + \xi^{(2)}\rho) Y^{(1)} + \xi^{(2)} \sqrt{1-\rho^2}\, Y^{(2)}.$$

Due to (8), regular variation and independence of $Y^{(i)}$ imply that

$$\begin{aligned} \gamma_\xi^* &= \lim_{t\to\infty} \frac{\mathrm{P}(\xi^\top X > t)}{\mathrm{P}(e_1^\top X > t)} \\ &= \lim_{t\to\infty} \frac{\mathrm{P}((\xi^{(1)} + \xi^{(2)}\rho) Y^{(1)} > t)}{\mathrm{P}(Y^{(1)} > 1)} + \lim_{t\to\infty} \frac{\mathrm{P}(\xi^{(2)}\sqrt{1-\rho^2}\, Y^{(2)} > t)}{\mathrm{P}(Y^{(1)} > 1)} \end{aligned}$$

(cf. Embrechts *et al.* 1997, Lemma 1.3.1). Moreover, from $Y^{(1)} \stackrel{d}{=} Y^{(2)}$ and $Y^{(2)} \in \mathcal{RV}_{-\alpha}$ we obtain that

$$\begin{aligned} \lim_{t\to\infty} \frac{\mathrm{P}\left(\xi^{(2)}\sqrt{1-\rho^2}\, Y^{(2)} > t\right)}{\mathrm{P}\left(X^{(1)} > t\right)} &= \lim_{t\to\infty} \frac{\mathrm{P}\left(Y^{(2)} > t / \left|\xi^{(2)}\sqrt{1-\rho^2}\right|\right)}{\mathrm{P}(X^{(2)} > t)} \\ &= \left|\xi^{(2)}\sqrt{1-\rho^2}\right|^\alpha. \end{aligned}$$

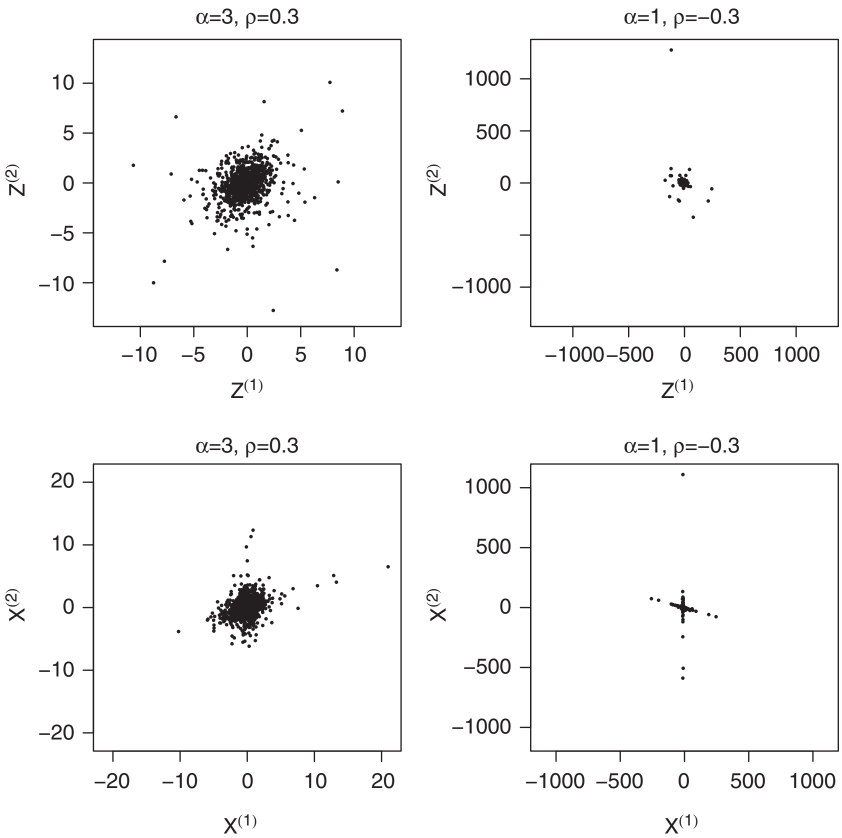

Figure 4. Elliptical vs. linear t model with a degrees of freedom and "correlation" parameter ρ: 1000 simulated i.i.d. observations of the random vector Z defined in (14) and the random vector $X = AY$ with $Y^{(i)}$ i.i.d. $\sim t(\alpha)$.

Similar arguments yield

$$\lim_{t\to\infty}\frac{\mathrm{P}((\xi^{(1)}+\xi^{(2)}\rho)\mathrm{Y}^{(1)}>t)}{\mathrm{P}(\mathrm{Y}^{(1)}>1)}=\lim_{t\to\infty}\frac{\mathrm{P}\big(\mathrm{Y}^{(1)}>t/|\xi^{(1)}+\xi^{(2)}\rho|\big)}{\mathrm{P}(\mathrm{Y}^{(1)}>t)}$$
$$=\left|\xi^{(1)}+\xi^{(2)}\rho\right|^{\alpha}.$$

Hence

$$(\gamma_{\xi}^{*}(X))^{1/\alpha}=\left(\left|\xi^{(1)}+\xi^{(2)}\rho\right|^{\alpha}+\left|\xi^{(2)}\sqrt{1-\rho^2}\right|^{\alpha}\right)^{1/\alpha}. \tag{15}$$

In contrast to the diversification effects in elliptical models, this expression depends on α. If $\rho \geq 0$, then the function $\xi \mapsto \gamma_{\xi}^{*1/\alpha}$ is convex for $\alpha \geq 0$ and concave for $\rho \leq 0$. In particular, $\gamma_{\xi}^{*1/\alpha}$ is linear in ξ for $\rho \geq 1$ and $\alpha = 1$. This means that VaR is asymptotically additive in this case:

$$\lim_{\lambda\nearrow 1}\frac{\mathrm{VaR}_{\lambda}(\xi^{\top}X)}{\xi^{(1)}\mathrm{VaR}_{\lambda}(X^{(1)})+\xi^{(2)}\mathrm{VaR}_{\lambda}(X^{(2)})}=1. \tag{16}$$

Figure 5 shows plots of the asymptotic risk profile $\gamma_{\xi}^{*1/\alpha}$ in the linear model for selected values of α and ρ. If $\rho < 0$, then the asymptotic portfolio risk is piecewise linear for $\alpha = 1$ and piecewise concave

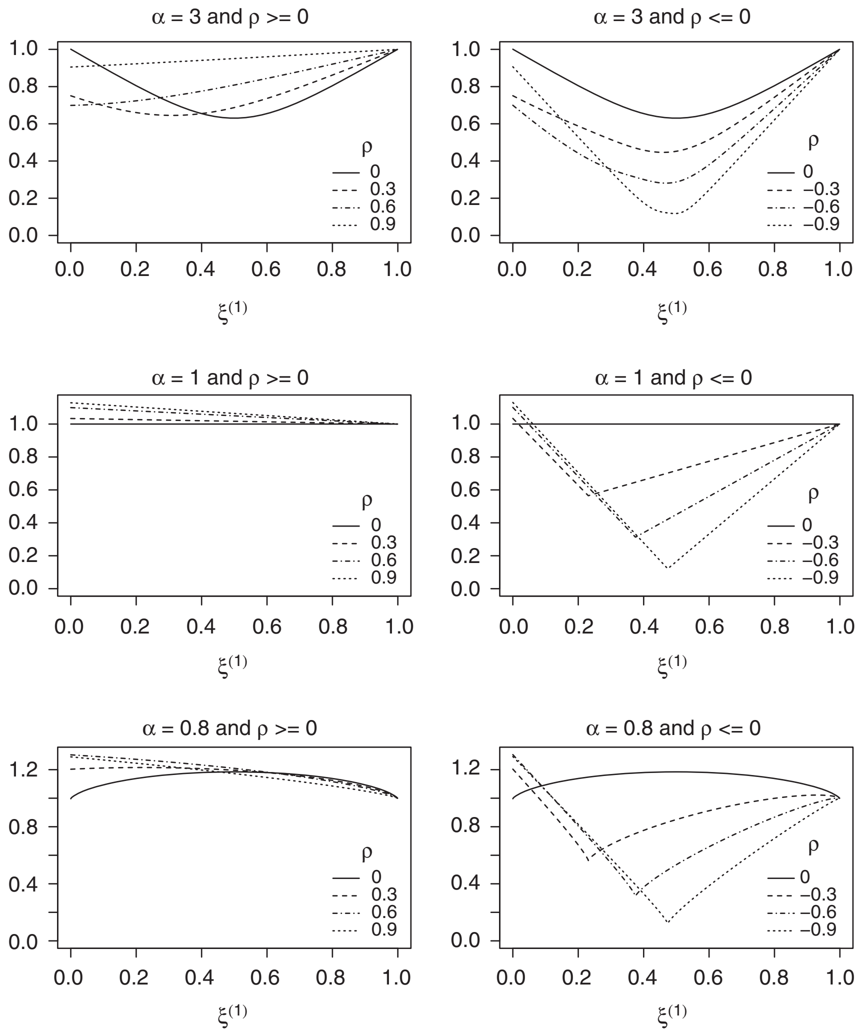

Figure 5. Asymptotic VaR ratio $\gamma_\xi^{*1/\alpha}$ in the linear heavy-tailed model.

for $\alpha < 1$. This demonstrates explicitly that the statements of Theorem 2.4(b, c) do not hold for X with $\Psi_X(\Sigma^d) < 1$.

This is extremely different from the elliptical t model, where the diversification effects are the same for all α (cf. Figure 2). Moreover, Figure 5 explains the violation of VaR subadditivity in the simulation study of Daníelsson *et al*. (2005), Daníelsson *et al*. (2012). The design of those simulation experiments was as follows. In each of $N = 1000$ repetitions, an i.i.d. sample of $X = AY$ with $Y^{(1)}$, $Y^{(2)}$ i.i.d. $\sim t$ was simulated (sample size $n = 10^6$). The empirical quantile of $X^{(1)} + X^{(2)}$ in each sample was used as the estimator of VaR_λ for $\lambda = 0.95$ and $\lambda = 0.99$. For $\alpha = 1$ and $\rho = 0$ or $\rho = 0.5$, the authors observed $\widehat{\text{VaR}}(X^{(1)} + X^{(2)}) > \widehat{\text{VaR}}(X^{(1)}) + \widehat{\text{VaR}}(X^{(2)})$ in roughly 50% of the $N = 1000$ repetitions. This simulation study was also repeated in Daníelsson *et al*. (2012) with smaller n and larger N. The results for smaller n are not as pronounced, but they still point to the same direction. For $\alpha = 1$, the subadditivity violation rates are above 40% and converge to 50% with increasing n, whereas the violation rates for $\alpha \geq 2$ become negligible for large n.

Looking at Figure 5 or at the formula (16), one recognizes immediately that VaR_λ is asymptotically linear for $\alpha = 1$ and $\rho \geq 0$. The convergence of the violation rate to 50% originates from the asymptotic normality of the estimator $\widehat{\mathrm{VaR}}$ for $n \to \infty$. With the large sample size $n=10^6$, the distribution of $\widehat{\mathrm{VaR}}$ is very close to normal with the true VaR as a mean. As the true VaR is almost additive, $\widehat{\mathrm{VaR}}(X^{(1)} + X^{(2)})$ lies below or above $\widehat{\mathrm{VaR}}(X^{(1)}) + \widehat{\mathrm{VaR}}(X^{(2)})$ with probability close to 1/2. Thus the 50% subadditivity violation rate is due to statistical noise, whereas the fact that we observe it for $\alpha = 1$ and $\rho \geq 0$ is a consequence of the asymptotic VaR additivity in this particular model.

The conclusion we can draw from the elliptical and the linear t examples presented above is that asymptotic diversification effects in the loss-gain case depend on both the tail index and on the dependence structure. Looking at samples and risk profiles on the entire portfolio range helps to discover such modelling traps.

3.4 Negative vs. positive dependence, short positions

We have already seen in Figure 5 that positive dependence ($\rho > 0$) and negative dependence ($\rho < 0$) have quite different implications on the diversification effects if loss-gain compensation is possible. This observation can be made more precise by adjusting the linear model $X = AY$ in a way that makes the asymptotic excess probabilities of $X^{(1)}$ and $X^{(2)}$ equal. From (15) we see that

$$\lim_{t\to\infty} \frac{\mathrm{P}(X^{(2)} > t)}{\mathrm{P}(X^{(1)} > t)} = \gamma^*_{e_2} = |\rho|^\alpha + \left(1-\rho^2\right)^{\alpha/2}.$$

Denoting $\widetilde{X}^{(2)} := b_\rho X^{(2)}$ with $b_\rho := {\gamma^*_{e_2}}^{-1/\alpha}$, we obtain that

$$\frac{\mathrm{P}(\widetilde{X}^{(2)} > t)}{\mathrm{P}(X^{(1)} > t)} = \frac{\mathrm{P}(X^{(2)} > t{\gamma^*_\xi}^{1/\alpha})}{\mathrm{P}(X^{(2)} > t)} \frac{\mathrm{P}(X^{(2)} > t)}{\mathrm{P}(X^{(1)} > t)} \to {\gamma^*_{e_2}}^{-1} \gamma^*_{e_2} = 1, \quad t \to \infty.$$

Hence, defining $\widetilde{X} = \widetilde{A} Y$ with $A = \begin{pmatrix} 1 & 0 \\ b_\rho \rho & b_\rho \sqrt{1-\rho^2} \end{pmatrix}$, we obtain

$$\gamma^*_{e_1}\left(\widetilde{X}\right) = \gamma^*_{e_2}\left(\widetilde{X}\right) = 1.$$

Moreover, analogously to (15), we obtain that

$$\left(\gamma^*_\xi(\widetilde{X})\right)^{1/\alpha} = \left(\left|\xi^{(1)} + \xi^{(2)} b_\rho \rho\right|^\alpha + \left|\xi^{(2)} b_\rho \sqrt{1-\rho^2}\right|^\alpha\right)^{1/\alpha}.$$

Figure 6 shows plots of $\left(\gamma^*_\xi(\tilde{X})\right)^{1/\alpha}$ for different α and ρ. The portfolio set also includes some short positions: $\xi^{(1)}$ ranges in $(-1, 2)$, and $\xi^{(2)} = 1 - \xi^{(1)}$. The corresponding plot for the elliptical model is shown in Figure 7.

First of all, we see that uniform ordering of portfolio risks is only possible for $\xi \in \Sigma^d$, i.e., if short positions are excluded. Furthermore, the re-weighting of $X^{(2)}$ to $\tilde{X}^{(2)}$ shows that the linear t model with $\rho \geq 0$ exhibits inversion of diversification effects for $\alpha < 1$ in the same way as it would be in the pure loss case. If, however, $\rho < 0$, then there is no inversion of this kind, but also no uniform ordering of portfolio risks even for $\xi \in \Sigma^2$.

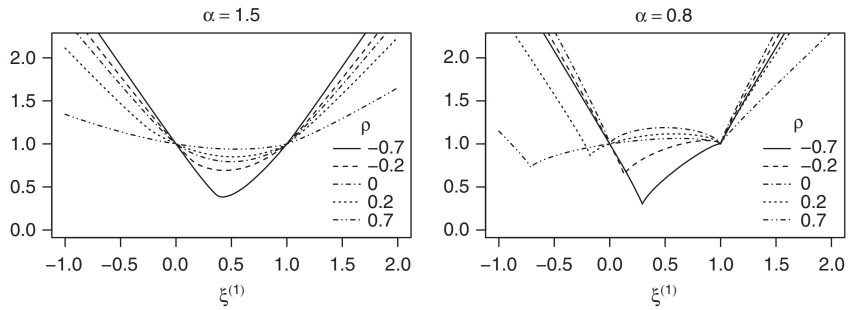

Figure 6. Asymptotic VaR ratio $\gamma_{\xi}^{*1/\alpha}$ in the linear model $\tilde{X} = \tilde{A}Y$ with balanced tails (short positions included).

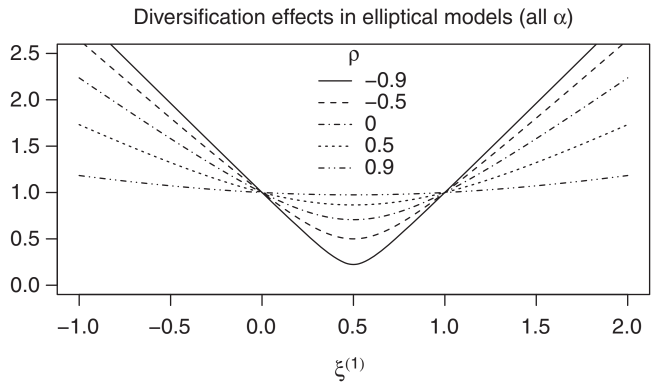

Figure 7. Asymptotic VaR ratio $\gamma_{\xi}^{*1/\alpha}$ in the elliptical model (short positions included).

Moreover, the right part of Figure 6 shows that the downward pinnacles of the portfolio risk we saw in Figure 5 also appear for $\rho > 0$. However, in this case they are associated with short positions. These pinnacles correspond to the optimal hedging strategies that would minimize the asymptotic portfolio risk. It is clear that in case of positive dependence, hedging needs a negative portfolio weight. There are no such effects in the elliptical model.

This comparison shows that in case of compensation between losses and gains, asymptotic sub- and superadditivity of VaR for $\alpha < 1$ is strongly influenced by the dependence structure. In particular, there is no general sub- or superadditivity result. Models that appear similar at a first glance can exhibit very different features. There seems to be a tendency of positive dependence to cause superadditivity. However, in models more complex than the linear model discussed here one would typically find a mixture of positive and negative dependence with an unclear outcome.

4 Conclusions and a word of warning

This paper demonstrates that VaR_λ in $\mathcal{MRV}$ models with tail index $\alpha \geq 1$ is asymptotically convex (and, in particular, subadditive) for $\lambda \nearrow 1$. On the other hand, VaR is asymptotically concave for random vectors in $\mathbb{R}_+^d$ that are $\mathcal{MRV}$ with $\alpha < 1$. Moreover, the influence of dependence on risk in the latter case is inverse, so that independence of asset returns increases the portfolio risk. For a more practical interpretation, see our discussion below Theorem 2.4.

We also demonstrate that general sub- or superadditivity results on VaR cannot be obtained for $\mathcal{MRV}$ models with $\alpha < 1$ if loss-gain compensation takes place in the tails. That is, with large losses and gains on the same scale, VaR can be asymptotically sub- or superadditive, or neither. The final result is determined by the particular tail dependence structure. A full description of the asymptotic diversification effects for VaR can be obtained from the asymptotic risk profile $\gamma_\xi^{*1/\alpha}$.

The examples discussed above also demonstrate the general fact that subadditivity of VaR *always* depends on the *combination* of marginal distributions and the dependence structure. In particular, the convexity results for $\alpha \geq 1$ do not hold if the assumption $X \in \mathcal{MRV}_{-\alpha,\Psi}$ is replaced by $X^{(i)} \in \mathcal{RV}_{-\alpha}$ for all i. That is, the influence of marginal distributions on the subadditivity of VaR cannot be clarified without assumptions on the dependence structure. All statements like "VaR is subadditive for Gaussian returns" or "VaR is subadditive for heavy-tailed returns with tail index $\alpha \geq 1$" are at least misleading. Strictly speaking, statements on VaR subadditivity that do not specify the dependence structure are simply wrong. Given marginal distributions $F_1, \ldots, F_d$, one can always construct a dependence structure such that VaR is either subadditive, additive, or superadditive (cf. Embrechts & Puccetti, 2010).

One also has to be careful when specifying the dependence structure. There are several alternative methods, such as linear or non-linear regression and copulas. Dependence models that are equivalent for some margins can be very different for others. In particular, the equivalence of elliptical and linear models is a special property of Gaussian margins. The comparison of the elliptical t model with a linear model based on t distributed factors shows how different the results can be in the general case. This example also demonstrates how misleading a blind trust in correlation parameters can be (see also Embrechts *et al*. 2002).

Similar issues arise in copula models. One should always be aware that transformation of marginal distributions has a crucial impact on the sample shape (cf. Balkema *et al*. 2010). In particular, the notion of elliptical copulas means only that these copulas are obtained from elliptical distributions. Combining such a copula with new margins, one can get highly non-elliptical samples. Gaussian copulas, for example, belong to the class of elliptical copulas. However, it is well known that they are asymptotically independent (cf. Sibuya, 1959). Endowing a Gaussian copula with t margins, one obtains a cross-shaped sample, similar to (but not the same as) the linear heavy-tailed model discussed in Subsection 3.3.

There are lots of modelling traps related to dependence. The best way to recognize and to circumvent them in practice is looking at the generated samples. For questions related to portfolio VaR asymptotics in $\mathcal{MRV}$ models one can also look at the risk profile $\xi \mapsto \gamma_\xi^{*1/\alpha}$ on the entire portfolio set. As demonstrated above, this kind of analysis helps to understand the difference between elliptical and linear models and also explains the simulation results of Daníelsson *et al*. (2005).

After various general and model specific conclusions outlined above, a word of warning should be said concerning the non-asymptotic sub- and superadditivity of VaR. The results presented here affect only the limit of VaR ratios, and the quality of the approximation they give to non-asymptotic problems depends on the distance between the true conditional excess distribution $\mathcal{L}((\|X\|, \|X\|^{-1}X) \mid \|X\| > t)$ and the limit distribution $\rho_\alpha \otimes \Psi$. The convergence rate of portfolio

VaR ratios depends on so-called second order parameters, and can be quite slow (cf. Degen *et al.* 2010). Moreover, the second-order parameters are even more difficult to estimate than the first-order parameters α and Ψ. This may put a practitioner into the paradoxical situation where the limit $\gamma_{\xi}^{*1/\alpha}$ is easier to estimate than the distance between this limit and the non-asymptotic quantity it should approximate. Being very useful from the qualitative point of view, VaR asymptotics should not be trusted blindly in quantitative non-asymptotic applications. No matter how fascinating theoretical results in a particular model may be, good practice always needs a humble approach and an appropriate awareness of the real questions and the relevant properties of the underlying data.

Acknowledgements

Georg Mainik thanks RiskLab at ETH Zürich for financial support. The authors would like to thank the anonymous referee for several constructive comments to an earlier version of the paper.

References

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**(3), 203–228.

Balkema, A.A., Embrechts, P. & Nolde, N. (2010). Meta densities and the shape of their sample clouds. *Journal of Multivariate Analysis*, **101**(7), 1738–1754.

Basrak, B., Davis, R.A. & Mikosch, T. (2002). A characterization of multivariate regular variation. *The Annals of Applied Probability*, **12**(3), 908–920.

Bingham, N.H., Goldie, C.M. & Teugels, J.L. (1987). *Regular Variation*. Cambridge University Press.

Boman, J. & Lindskog, F. (2009). Support theorems for the Radon transform and Cramér-Wold theorems. *Journal of Theoretical Probability*, **22**(3), 683–710.

Daníelsson, J., Jorgensen, B.N., Samorodnitsky, G., Sarma, M. & Vries, C.G. de (2005). *Subadditivity re-examined: the case for Value-at-Risk*. FMG Discussion Papers. London School of Economics.

Daníelsson, J., Jorgensen, B.N., Samorodnitsky, G., Sarma, M. & Vries, C.G. de (2012). Fat Tails, VaR and Subadditivity. *Preprint*.

Degen, M., Lambrigger, D.D. & Segers, J. (2010). Risk concentration and diversification: Second-order properties. *Insurance: Mathematics and Economics*, **46**(3), 541–546.

Delbaen, F. (2009). Risk measures for non-integrable random variables. *Mathematical Finance*, **19**(2), 329–333.

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997). *Modelling Extremal Events*. Springer.

Embrechts, P., Lambrigger, D.D. & Wüthrich, M.V. (2009). Multivariate extremes and the aggregation of dependent risks: examples and counterexamples. *Extremes*, **12**(2), 107–127.

Embrechts, P., McNeil, A.J. & Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In: *Risk Management: Value at Risk and Beyond* Ed. by M. A. H. Dempster. Cambridge University Press.

Embrechts, P. & Puccetti, G. (2010). Risk Aggregation. In: *Copula Theory and Its Applications*. Ed. by P. Jaworski, F. Durante, W.K. Härdle & T. Rychlik. Springer.

Fama, E.F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, **38**(1), 34–105.

Fama, E.F. & Miller, M.H. (1972). *The theory of finance*. Holt, Rinehart, and Winston, New York.

Haan, L. de & Ferreira, A. (2006). *Extreme Value Theory*. Springer.

Hardy, G.H., Littlewood, J.E. & Pólya, G. (1934). *Inequalities*. Cambridge University Press.

Hofert, M. & Wüthrich, M. (2011). Statistical Review of Nuclear Power Accidents. *Preprint*.

Hult, H. & Lindskog, F. (2002). Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied Probability*, **34**(3), 587–608.
Hult, H. & Lindskog, F. (2006). Regular variation for measures on metric spaces. Institut Mathématique. Publications. *Nouvelle Série*, **80**(94), 121–140.
Ibragimov, R., Jaffee, D. & Walden, J. (2009). Nondiversification Traps in Catastrophe Insurance Markets. *Review of Financial Studies*, **22**(3), 959–993.
Ibragimov, R., Jaffee, D. & Walden, J. (2011). Diversification disasters. *Journal of Financial Economics*, **99**(2), 333–348.
Mainik, G. *On Asymptotic Diversification Effects for Heavy-Tailed Risks*. PhD thesis. University of Freiburg, 2010.
Mainik, G. & Rüschendorf, L. (2010). On optimal portfolio diversification with respect to extreme risks. *Finance and Stochastics*, **14**(4), 593–623.
Mainik, G. & Rüschendorf, L. (2012). Ordering of multivariate probability distributions with respect to extreme portfolio losses. *Statistics & Risk Modeling*, **29**(1), 73–106.
Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *The Journal of Business*, **36**(4), 394–419.
Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, **7**(1), 77–91.
McNeil, A.J., Frey, R. & Embrechts, P. (2005). *Quantitative Risk Management*. Princeton University Press.
Mikosch, T. (2003). Modeling Dependence and Tails of Financial Time Series. In: *Extreme Values in Finance, Telecommunications, and the Environment*. Chapman and Hall/CRC, pp. 185–286.
Moscadelli, M. (July 2004). *The modelling of operational risk: experience with the analysis of the data collected by the Basel Committee*. Temi di discussione (Economic working papers) 517. Bank of Italy, Economic Research Department.
Resnick, S.I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer.
Resnick, S.I. (2007). *Heavy-Tail Phenomena*. Springer.
Rootzén, H. & Klüppelberg, C. (1999). A Single Number Can't Hedge against Economic Catastrophes. *Ambio*, **28**(6), 550–555.
Sibuya, M. (1959). Bivariate extreme statistics, I. *Annals of the Institute of Statistical Mathematics*, **11**(2). 10.1007/BF01682329, 195–210.
Zhou, C. (2010). Dependence structure of risk factors and diversification effects. *Insurance: Mathematics and Economics*, **46**(3), 531–540.

A: Spectral measures in linear models

In this section we derive an explicit representation for the 4-point spectral measure of the linear heavy-tailed model. Independence of $Y^{(1)}$ and $Y^{(2)}$ implies that

$$\frac{\bar{F}_{\|Y\|_1}(r)}{2\bar{F}_{|Y^{(1)}|}(r)} = \frac{\mathrm{P}\{|Y^{(1)}| + |Y^{(2)}| > r\}}{\mathrm{P}\{|Y^{(1)}| > r\} + \mathrm{P}\{|Y^{(2)}| > r\}} \to 1$$

as $y \to \infty$ (cf. Embrechts *et al.* 1997, Lemma 1.3.1). For any fixed $\varepsilon > 0$ this yields

$$\mathrm{P}\{|Y^{(1)}| > \varepsilon r, |Y^{(2)}| > \varepsilon r | \|Y\|_1 > r\} \leq \frac{(\bar{F}_{|Y^{(1)}|}(\varepsilon r))^2}{\bar{F}_{\|Y\|_1}(r)} \to 0, \quad r \to \infty.$$

This implies that observations of Y with large $\|Y\|_1$ are concentrated around the coordinate axes. By symmetry arguments we easily obtain that the conditional probabilities $\mathrm{P}\{\pm Y^{(i)} \leq \varepsilon r | \|Y\|_1 > r\}$ are equal for $i = 1, 2$. This entails that $Y \in \mathcal{MRV}_{\alpha,\Psi_Y}$ with spectral measure

$$\Psi_Y = \frac{1}{4}(\delta_{v_1} + \delta_{v_2} + \delta_{v_3} + \delta_{v_4}),$$

where $v_1 := (1,0)^\top$, $v_2 := (1,0)^\top$, $v_3 := -v_1$, and $v_4 := -v_2$. The corresponding exponent measure $\nu_Y = c(\rho_\alpha \otimes \Psi_Y) \circ \tau$ (cf. (4) and (5)) is concentrated on the coordinate axes. As the constant $c > 0$ can be chosen freely, we simplify the writing by setting $c := 4$. In this case we have that

$$\nu_Y(\pm(1,\infty)e_i) = 1, \quad i = 1.2,$$

with the notation $Bv := \{x \in \mathbb{R}^2 : x = rv, r \in B\}$ for $v \in \mathbb{R}^2$ and $B \subset \mathbb{R}$.

To obtain the spectral measure Ψ_X of the linear model $X = AY$ with the matrix A defined in (13), consider the definition (4) of $\mathcal{MRV}$. Denote $T_A(x) := Ax$ for $x \in \mathbb{R}^d$. Since A is invertible, we have $T_A^{-1} = T_{A^{-1}}$. The inverse of A is given by

$$A^{-1} = \begin{pmatrix} 1 & 0 \\ -\rho/\sqrt{1-\rho^2} & 1/\sqrt{1-\rho^2} \end{pmatrix}.$$

As Y satisfies (4) with $\nu = \nu_Y$ and some sequence a_n, we immediately obtain

$$n\mathrm{P}\{a_n^{-1}X \in B\} = n\mathrm{P}\{a_n^{-1}Y \in T_A^{-1}(B)\} \to \nu_Y(T_A^{-1}(B))$$

for compact $B \in \mathcal{B}([-\infty,\infty]^d \backslash \{0\})$. That is, X satisfies (4) with the same a_n and $\nu = \nu_X := \nu_Y \circ T_A^{-1}$. In particular, the support of ν_X is the support of ν_Y transformed by T_A:

$$\mathrm{supp}(\nu_X) = T_A(\mathrm{supp}(\nu_Y)) = T_A\left(\bigcup_{i=1}^{4}(0,\infty)v_i\right) = \bigcup_{i=1}^{4}(0,\infty)Av_i.$$

Hence the support of the spectral measure Ψ_X on $\mathbb{S}_1^2$ is given by

$$\mathrm{supp}(\Psi_X) = \mathrm{supp}(\nu_X) \cap \mathbb{S}_1^2 = \{w_1, w_2, w_3, w_4\}$$

where $w_i := \|Av_i\|_1^{-1}Av_i$. That is, $w_1 = (1 + |\rho|)^{-1}(1,\rho)^\top$, $w_2 = (0,1)^\top$, $w_3 = -w_1$, $w_4 = -w_2$. The spectral measure Ψ_X is given by

$$\Psi_X(\{w_i\}) = \frac{\nu_X((1,\infty)w_i)}{\sum_{j=1}^{4}\nu_X((1,\infty)w_j)}.$$

Since $\nu_X((1,\infty)w_i) = \nu_Y((1,\infty)A^{-1}w_i) = \|A^{-1}w_i\|_1^{-\alpha}$ for $i = 1, \ldots, 4$, we obtain that

$$\Psi_X(\{w_1\}) = \Psi_X(\{w_3\}) = \frac{(1 + |\rho|)^\alpha}{2(1 + |\rho|)^\alpha + 2(1-\rho^2)^{\alpha/2}},$$

$$\Psi_X(\{w_2\}) = \Psi_X(\{w_2\}) = \frac{(1-\rho^2)^{\alpha/2}}{2(1 + |\rho|)^\alpha + 2(1-\rho^2)^{\alpha/2}}.$$

□